EPA Technical Team Level of Effort (LOE) for Investigations Designed to Evaluate Risks of Contaminants to Benthic Invertebrate Communities in the Upper Columbia River (Sediment Toxicity LOE)

January 13th, 2010
Developed by the USEPA Technical Team (David Charters, Steve Cox, Bruce Duncan, Marc Greenberg, Chris Ingersoll, Cameron Irvine, Don MacDonald, Patrick Moran, Dave Mount, and Jeff Steevens), with consideration of comments provided by the Participating Parties

1.0 Goals

The purpose of this document is for the USEPA Technical Team to provide a summary of the level of effort expected for investigations designed to evaluate the risks to benthic invertebrate communities exposed to chemicals of potential concern (COPCs; which are also referred to as chemicals of interest or COIs) in the Upper Columbia River (UCR). Specifically, the goal of this sediment sampling component of the baseline ecological risk assessment (BERA) is to evaluate risks of metals and other chemicals [as identified by the screening-level ecological risk assessment (SLERA), for the site] of interest to sediment-dwelling invertebrates inhabiting the UCR. To achieve this goal, site-specific relationships between sediment chemistry and sediment toxicity will be established to evaluate the toxicity associated with exposure to sediment and/or pore-water in the study area. A variety of exposure metrics (e.g., total metals, SEM-AVS, SEM-AVS/f_{oc}, BLM parameters) should be considered in this evaluation to maximize the potential for effectively addressing the bioavailability of site COPCs.

Once these relationships are developed, there will be a need to evaluate potential effects throughout the study area. To that end, future sediment sampling will focus on evaluating the nature and extent of unacceptable risk from contaminants, including transport and fate of COPCs as applicable. Limited site-specific sediment toxicity data are currently available for the site and the data that are available are dated. This will necessitate collection of additional data on the distribution of the selected indicators of sediment toxicity throughout the UCR as part of a future sampling effort.

2.0 Evaluating Causes of Observed Toxicity

A key objective of the sediment toxicity evaluation is to establish relationships between sediment characteristics (i.e., sediment chemistry, pore-water chemistry, grain size, TOC, and other metrics) and toxicity to benthic organisms, as measured in laboratory toxicity tests. Toxicity testing should include the five principal toxicity tests described in Section 6.0. Survival, growth, and reproduction of benthic invertebrates need to be evaluated on all sediment samples to provide synoptic information relevant to the selected assessment endpoint for this portion of the BERA. A sub-set of the highly contaminated samples should be further tested to determine if the shorter term tests provide a reasonable basis for estimating reproductive effects in midge and/or amphipods in UCR sediment (see Section 6.0 for further information).

While correlations may be found between sediment toxicity and many sediment characteristics, the evaluation of risks to benthic invertebrates will better inform remedial actions if the specific cause(s) of sediment toxicity can be determined. The scope of study necessary to fulfill this objective is dependent on the strength and significance of the relationships between sediment toxicity test results and the concurrent measures of exposure (i.e., sediment and pore-water chemistry). If toxicity can clearly be attributed to one or more COPC, based on measured exposure/bioavailability and known potency to tested species, then further evaluation may not be necessary. However, if these causal relationships are not reasonably identifiable, then supplemental studies will likely be needed. Additional lines of evidence could include collection of water only toxicity data with COIs (or mixtures thereof) to determine exposure-response relationships and/or confirm the applicability of toxicity models (e.g., biotic ligand model [BLM]). Another tool that could be used in this regard is toxicity identification evaluation (TIE) procedures, as means to confirm the role of cationic metals (if they are suspected causal toxicants) in producing observed toxicity. This latter approach may be particularly useful if EPA concludes there are outstanding questions as to the relative importance of sediment contaminants versus potential confounding factors. Tissue residue data may also be useful in understanding causality if appropriate tissue residue-response relationships can be identified.

3.0 Lines of Evidence

Evaluation of risks to the benthic invertebrate community should utilize multiple lines of evidence, including:

- Whole-Sediment Chemistry;
- Whole-Sediment Toxicity;
- Pore-Water Chemistry;
- Benthic Invertebrate Tissue Chemistry; and,
- Surface-Water Chemistry

The assessment of risks to benthic invertebrates should initially evaluate each of these lines of evidence independently. Subsequently, multiple lines of evidence may be integrated to develop a weight-of-evidence for assessing risks to this receptor group. It is anticipated that TECK may propose such a weight-of-evidence assessment framework for review by USEPA and the Participating Parties.

4.0 Chemicals of Potential Concern

Upon approval, the draft SLERA will provide a basis for identifying COPCs that need to be addressed in the BERA. Any COPC refinement must be conducted in accordance with USEPA guidance. Accordingly, COPCs should not be eliminated without appropriate rationale. A preliminary list of COPCs to be measured in sediments, pore water, and invertebrate tissues is provided in Table 1.

5.0 Sediment and Slag Characterization

Data describing the sediment used in toxicity tests will be potentially useful for explaining concentration-response relationships. If toxicity is observed, then all sediment samples should be described to help identify characteristics that are related to toxicity. Separate sub-samples should be collected at the time of sampling for each of these potential measurements, and archived for possible analysis. There is currently no accepted method for characterizing slag in sediment samples; therefore, multiple measurements should be made to describe the sediment and slag therein. These measurements may provide insight into toxicity patterns at the UCR. These should not be considered definitive measurements of slag or slag content, but are intended to provide possible data that may help to explain where and under what conditions toxicity is observed.

- Grain size (including % sand/silt/clay) should be measured in all sediment samples. Although grain size is not equivalent to slag, this measurement has shown a potential relationship with slag content (July 2009 draft BERA work plan).

- A qualitative estimate of occurrence of slag can also be made for all samples with zinc concentrations greater than 1000 mg/kg. The estimate is based on the concentrations of a suite of elements that are enriched in slag compared to geologic materials common to the area. From the limited data that are available, a suite of slag-enriched elements includes, but is not limited to, zinc, copper, antimony, iron, and calcium. Those limited data show typical slag from the Trail smelter has trace-element concentrations exceeding 20,000 mg/kg zinc, 3,000 mg/kg copper, 100 mg/kg antimony, 300,000 mg/kg iron, and 90,000 mg/kg calcium. If the relative concentration (in percent) measured in a sediment sample of five enrichment indicator trace-elements are similar, then the median relative concentration of the five enrichment indicator trace-elements could be used as a qualitative estimate of the occurrence of slag. This estimation method should not be used below about 5 or 10 percent, because all of these elements are present in various geological materials. Prior to field work, this estimation method should be refined and tested using additional data on the trace element composition of slag and tested with samples of native sediment spiked with known quantities of slag.

- A more quantitative measure of slag content based on petrographic analysis could also be conducted on selected samples demonstrating a range of toxicity levels. Petrographic analyses may include:

    - Quantitative measurement involving individual point counts of sediment grains (minimum 200 grains/sample) for the characteristic grain size of the sediment sample and visible slag grains.

    - X-ray fluorescence (XRF) or x-ray absorption spectrometry (using a synchrotron) to describe the shape and porosity of sediment grains.

6.0 Whole-Sediment Chemistry and Whole-Sediment Toxicity

Matching whole-sediment toxicity and whole-sediment chemistry data need to be collected to support the development of site-specific concentration-response relationships for the study area, as described in Section 1. As the key assessment endpoint for this portion of the BERA is the survival, growth, and reproduction of benthic invertebrates, there is a need to select a suite of laboratory toxicity tests that will provide the data and information needed to evaluate the status of the benthic invertebrate community. The tests that need to be conducted on all sediment samples include:

- 10-d whole-sediment toxicity tests with the midge, *Chironomus dilutus* (Endpoints - survival, weight, and biomass; USEPA 2000, ASTM 2009a);
- 28-d whole-sediment toxicity tests with the amphipod, *Hyalella azteca* (Endpoints - survival, weight, and biomass; USEPA 2000, ASTM 2009a); and,
- 28-d whole-sediment toxicity tests with the mussel, *Lampsilis siliquoidea* (Endpoints - survival, weight, and biomass; USEPA 2000, ASTM 2009ab).

This suite of toxicity tests will provide useful information for evaluating the survival and growth of benthic invertebrates; however, these toxicity tests do not address the potential for effects on reproduction. For this reason, a limited sub-sample (minimum of 12) of sediment samples with moderately to highly elevated metal concentrations should be selected for conducting long-term toxicity tests for estimating reproductive effects. More specifically, the following toxicity tests, evaluating reproduction in midges and amphipods, should be conducted on the selected sediment sub-samples:

- 53-d whole-sediment toxicity tests with the midge, *Chironomus dilutus* (Endpoints - survival, weight, biomass, emergence, eggs/surviving female, egg hatching, viability of young; using the adapted method starting with 7-d-old larvae; USEPA 2000, ASTM 2009a); and,
- 42-d whole-sediment toxicity tests with the amphipod, *Hyalella azteca* (Endpoints - survival, weight, biomass, neonates/surviving female; USEPA 2000, ASTM 2009a).

The toxicity testing laboratories must demonstrate acceptable control results with multiple control materials, including their usual control material, West Bearskin Lake, and Browns Lake. Importantly, laboratories conducting sediment toxicity tests must provide historical data demonstrating that their control sediments and overlying water provide acceptable conditions for conducting these tests. Specifically, see the requirements outlined in Appendix 1 for documenting the capability of toxicity testing laboratories to conduct sediment toxicity tests with amphipods, midge, and mussels based on guidance provided in USEPA (2000) and in ASTM (2009a,b,c,d).

7.0 Pore-Water Chemistry

For all sediment samples that are collected to support the development of concentration-response relationships, pore-water chemistry needs to be measured. Pore-water quality should be monitored at the start of the toxicity or bioaccumulation tests (e.g., Day -7 when sediments placed in exposure beakers: pH, eH, ORP, ammonia, dissolved organic carbon, dissolved oxygen, free sulfide, hardness, alkalinity, conductivity, major cations, major anions; Table 1). Sediments should be equilibrated for about 7 days before the start of the toxicity or bioaccumulation tests (Ingersoll *et al*. 2008). Pore-water samples should also be sampled during toxicity tests by inserting peepers (for seven days) into additional chemistry replicates that are loaded with animals, sediment, and food (Ingersoll *et al.* 2008). The selected pore-water chemistry measurements should provide the data needed to apply the biotic ligand model to evaluate metal toxicity in sediment pore water. The water in the peepers should be scanned for a suite of metals of interest (Table 1).

8.0 Holding Times for Toxicity Tests

USEPA (2000) and ASTM (2009a) guidance on holding times should be followed for sediment samples that will be used in toxicity testing, if the results are going to be used to make statements about the toxicity on a spatial scale. For TIE, holding times are less important because the samples will be retested to confirm that they are still toxic before the TIE manipulations are initiated (USEPA 2007).

9.0 Bioaccumulation and Benthic Invertebrate-Tissue Chemistry

Invertebrate-tissue chemistry represents an important line of evidence for evaluating risks to benthic invertebrates associated with exposure to COPCs in the study area. More specifically, data from laboratory bioaccumulation tests provide essential information on the bioavailability of sediment-associated COPCs and on their accumulation in invertebrate tissues. Matching tissue chemistry and sediment toxicity data from laboratory toxicity tests can provide the information needed to identify critical body residues of COPCs. In turn, this information can be used to interpret field collected invertebrate-tissue chemistry data.

Bioaccumulation of sediment-associated COPCs should be evaluated in 28-d whole-sediment bioaccumulation tests with the oligochaete, *Lumbriculus variegatus*, using splits of 15 to 20 sediment samples collected to support toxicity testing. Test methods should follow appropriate guidance (e.g., USEPA 2000 and ASTM 2009d). These sediment samples should be selected to provide a large concentration gradient of metals in sediments that are considered to be slag-dominated. A preliminary list of COIs for tissue residues are listed in Table 1.

Additionally, tissue residues (i.e., body burdens) should be measured in midge exposed to the same 15 or 20 samples from the site that are used for oligochaete bioaccumulation testing. Test methods should follow those described in USEPA (2000). Midge, for analysis of tissue chemistry, should be obtained from additional replicates of samples used for the 10-day midge

bioassay (i.e., exposures for tissue chemistry should be conducted concurrently with the toxicity tests). The objective of this study element is to develop site-specific relationships between toxicity and body burdens of COIs that can be used to evaluate the data on invertebrate-tissue chemistry for tissue samples collected from hard bottom substrates and other locations.

Thirdly, the concentrations of COIs in field-collected invertebrate-tissue samples should be assessed in the riverine reach to evaluate exposures in organisms with different exposure pathways than those for benthic invertebrates evaluated in the sediment toxicity tests. This collection effort would target two or three specific taxa that are sufficiently abundant across all, or most, riverine sites such that a relative tissue accumulation level could be interpreted. The sampling effort should include collection of invertebrate-tissue samples at least 25 locations (not including reference samples) over a demonstrated contaminant gradient. Some taxonomic expertise will be needed on-site at the time of collection to properly match taxa across sites. Replication at each site (i.e., multiple samples of the same species at a given location) would depend upon organism abundance following an equal collection effort, but should be sought to help evaluate inter-site variability. Collected invertebrates should be depurated (if possible), rinsed, and pooled from each sampling location to achieve minimum mass requirements. Given the low sample mass expected from this taxa-specific tissue comparison, an ICP/MS analytical method with its minimal mass requirements should be used. Tissue concentrations can be compared between exposure sites and reference sites and to literature-derived critical body residues. Depending upon the outcome of this study, further investigation of invertebrate metal accumulation and/or toxicity may be warranted.

Finally, traditional benthic community collection of invertebrates that may take place at these sites would help infer the 'representativeness' of the particular taxa chosen for the tissue accumulation work described above.

10.0 Toxicity Identification Evaluations

TIE-type manipulations may be particularly useful if there are questions as to the relative importance of sediment contaminants versus potential confounding factors (USEPA 2007). These manipulations could include addition of cation exchange resin, addition of resin without cation exchange capacity, or supplementing acid-volatile sulfide, as means to confirm the role of cationic metals (if they are suspected causal toxicants in producing observed toxicity). Application of other manipulations may be warranted as indicated by existing guidance and other data collected as part of the overall sediment LOE.

11.0 Selection of Sediment Sampling Locations

Experience at other sites indicates that 100 to 130 samples should be collected to provide sufficient data to support the derivation of concentration-response relationships for geographic areas comparable to the UCR. Sampling locations should be selected to provide a strong concentration gradient for metals, from those that are reflective of reference conditions (i.e.,

mean PEC-Q of <0.1) to those that include high concentrations of trace metals and/or other COPCs, percentages of slag. It may also be useful to ensure that these sampling locations encompass the a range of potential slag content, based on qualitative estimates of the occurrence of slag described in Section 5 or on other methods (i.e., slag-dominated, non-slag dominated, and other). Other parameters may be identified that have associations with sediment toxicity; therefore, if there are multiple sampling events needed to collect the target number of samples, then the sediment toxicity sampling program should be flexible enough to allow changes to the sampling design that incorporate new information as it identified.

It is understood that sediments within the UCR exhibit a range of characteristics. For this reason, the sampling program should be designed to provide broad spatial coverage and consider such factors as distance from source, water depth, water velocity, and others. Samples should include some targeted sediment collected from under cobble and interstitial spaces between cobbles. This is particularly pertinent since large deposits of slag have been found trapped under and between cobbles. These interstitial spaces serve as important habitat for aquatic organisms, and no evaluations have been done that would determine whether the slag found under cobbles is different from slag found in other substrates.

A reference envelope approach should be used to evaluate the sediment toxicity generated under this sampling program. USEPA can provide specific guidance on the interpretation of the sediment toxicity data using the reference envelope approach at a later date. The sampling program should target acquisition of matching sediment chemistry and sediment toxicity data at about 15 reference locations in the study area (i.e., within site and/or outside site reference locations, with at least two-thirds of these located outside the influence of releases from the facility).

12.0 Sample Collection Methods

Methods for collection, handling, and storage of sediments should be in general concordance with guidance provided in ASTM (2009c). Because this study is primarily designed to evaluate relationships between sediment chemistry and toxicity and because chemical analysis must target the <2.00 mm fraction, it is anticipated that all sediment samples will be sieved in the field to <2.00 mm. Following sieving, samples should be homogenized and split to support chemical analysis, toxicity testing, and bioaccumulation testing. In general, measures of *in situ* pore-water metals, and BLM parameters should be made with peepers, along with measures of *in situ* SEM-AVS/f_{oc} in site sediments. Likewise, pore-water measurements and sediment SEM-AVS/f_{oc} should also be made in chemistry-only replicate test beakers at the end of each toxicity and bioaccumulation test. Comparisons between *in situ* and laboratory measurements would evaluate potential differences in sediment conditions from collection, handling, and storage to testing of sediments (pore-water conditions in tested sediments are described in Section 7 and Table 1). Chemistry should be re-evaluated in sediment samples stored for longer than the recommended holding times. *In situ* measures of pH, eH, and ORP may also be helpful in characterizing the sediments.

13.0 References Cited

American Society for Testing and Materials (ASTM). 2009a. Standard test method for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates (E1706-05). In ASTM Annual Book of Standards, Vol. 11.06, West Conshohocken, Pennsylvania.

American Society for Testing and Materials (ASTM). 2009b. Standard guide for conducting laboratory toxicity tests with freshwater mussels (E2455-06). In ASTM Annual Book of Standards, Vol. 11.06, West Conshohocken, Pennsylvania.

American Society for Testing and Materials (ASTM). 2009c. Standard guide for collection, storage, characterization, and manipulation of sediments for toxicological testing and for selection of samplers used to collect benthic invertebrates (E1391-03 (2008)). In ASTM Annual Book of Standards, Vol. 11.06, West Conshohocken, Pennsylvania.

American Society for Testing and Materials (ASTM). 2009d. Standard guide for conducting bioaccumulation of sediment-associated contaminants by benthic invertebrates (E1688-00a. In ASTM Annual Book of Standards, Vol. 11.06, West Conshohocken, Pennsylvania.

Cain DJ, Luoma SN, Wallace WG. 2003. Linking metal bioaccumulation of aquatic insects to their distribution pattersn in a mining-impacted river. *Environ Toxicol Chem* 23:1463-1473.

Clements WH, Carlisle DM, Courtney LA, Harray EA. 2002. Integrating observational and experimental approaches to demonstrate causation in stream biomonitoring studies. Environ Toxicol Chem 21:1138-1146.

Day KE, Dutka BJ, Kwan KK, Batista N, Reynoldson TB, Metcalfe-Smith JL. Correlations between solid-phase microbial screening assays, whole-sediment toxicity tests with macroinvertebrates and in situ benthic community structure. *J Great Lakes Res* 21:192-206.

Day KE, Maguire RJ, Milani D, Batchelor SP. 1998. Toxicity of tributyltin to four species of freshwater benthic invertebrates using spiked sediment bioassays. *Water Qual Res J Canada* 33:111-132.

Ingersoll CG, Wang N, Hayward JMR, Jones JR, Jones S. 2005. A field assessment of long-term laboratory sediment toxicity tests with the amphipod *Hyalella azteca. Environ Toxicol Chem* 24:2853-2870.

Ingersoll CG, Besser JM, Brumbaugh WG, Ivey CD, Kemble NE, Kunz JL, May TW, Wang N, MacDonald DD, Smorong DE. 2008. Sediment chemistry, toxicity, and bioaccumulation data report for the US Environmental Protection Agency – Department of the Interior sampling of metal-contaminated sediment in the Tri-state Mining District in Missouri, Oklahoma, and Kansas. Prepared by USGS, Columbia MO and MacDonald Environmental Sciences Ltd.,

Nanaimo, BC for the USEPA Kansas City, MO; USEPA, Dallas, TX; and USFWS, Columbia, MO.

Irving EC, Baird DJ, Culp JM. 2003. Ecotoxicological responses of the mayfly *Baetis tricaudatus* to dietary and waterborne cadmium: implications for toxicity testing. *Environ Toxicol Chem* 22:1058-1064.

U.S. Environmental Protection Agency. 2000. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates, second edition. EPA 823-B-99-007, Duluth, MN and Washington, DC.

U.S. Environmental Protection Agency. 2007. Sediment toxicity identification evaluation (TIE): Phases I, II, and III Guidance Document. EPA/600/R-07/08. , Washington, DC.

Wang N, Ingersoll CG, Greer IE, Hardesty DK, Ivey CD, Kunz JL, Brumbaugh WG, Dwyer FJ, Roberts AD, Augspurger T, Kane CM, Neves RJ, Barnhart MC. 2007. Chronic toxicity of copper and ammonia to juvenile freshwater mussels (Unionidae). *Environ Toxicol Chem* 26:2048-2056.

Table 1. Parameters of Interest for sediment toxicity characterization investigations.

Parameters	Field Sediment	Laboratory Bioassay Sediment	In Situ Pore-Water	Laboratory Bioassay Pore-Water (Day7)	Tissues
TAL Metals: Aluminum, Antimony, Arsenic, Barium, Beryllium, Cadmium, Chromium, Cobalt, Copper, Iron, Lead, Magnesium, Manganese, Mercury, Molybdenum, Nickel, Selenium, Silver, Thallium, Uranium, Vanadium, Zinc	x	x	x	x	x
Tin,	x	x			
Calcium, Potassium, Sodium	x	x			
Methyl Mercury	x	x	x	x	x
DDE/DDD/DDT	x	x	x	x	
Methoxychlor					
Dioxins/Furans	x	x	x	x	
PBDEs	x	x	x	x	
AVS	x	x			
SEM	x	x			
f_{oc}	x	x			
pH, eH, ORP,	x	x	x	x	
ammonia, dissolved organic carbon, dissolved oxygen, free sulfide, hardness, alkalinity, conductivity, major cations, major anions (BLM Parameters)			x	x	

Appendix 1 - Demonstrating Acceptable Laboratory Performance in (1) Whole-sediment Toxicity Tests Conducted with the Amphipod *Hyalella azteca,* with the Midge *Chironomus dilutus* or with Other Sediment Toxicity Testing Organisms (Mussels, Mayflies) and (2) Whole-sediment Bioaccumulation Tests Conducted with the Oligochaete *Lumbriculus variegatus*

This appendix summarizes ASTM (2009a) and USEPA (2000) recommendations that a testing laboratory should follow in order to demonstrate acceptable performance in whole-sediment laboratory toxicity tests conducted with the amphipod *Hyalella azteca* or with the midge *Chironomus dilutus*.

Guidance is summarized in Tables 1 to 5 regarding test acceptability requirements outlined in ASTM (2009a) and in USEPA (2000) for conducting long-term sediment toxicity tests with *H. azteca and C. dilutus.* Analogous tables summarizing test acceptability requirements for 10-d tests with these two species are provided in ASTM (2009a) and in USEPA (2000).

Guidance is also summarized in Tables 1 to 5 for conducting sediment toxicity tests with mussels and mayflies and in Tables 6 to 8 for conducting sediment bioaccumulation tests with the oligochaete *Lumbriculus variegatus*.

In addition to these test acceptability requirements, the following sections highlights key ASTM (2009a) and in USEPA (2000) recommendations that should be followed by a laboratory in order to demonstrate acceptable performance of sediment toxicity tests conducted with *H. azteca* or *C. dilutus* or with other sediment toxicity test organisms or bioaccumulation test organisms.

Summary of recommendations in ASTM (2009a) and in USEPA (2000) regarding requirements for a laboratory to demonstrate acceptable performance in toxicity tests conducted with the amphipod *Hyalella azteca* or with the midge *Chironomus dilutus* or sediment bioaccumulation tests with the oligochaete *Lumbriculus variegatus*. These recommendations are also generally applicable for testing with other organisms including mussels and mayflies

1. Tests should meet test acceptability requirements outlined in Tables 1 to 8 (summarized below from ASTM 2009a and USEPA 2000).
2. Overlying water should be a source of water that has been demonstrated to support survival, growth, and reproduction of *H. azteca* (Section A6.2.2 in ASTM 2009a and Section 14.2.2 in USEPA 2000). Variable success has been reported using reconstituted waters to culture or test *H. azteca* (Section 12.2.5.1 in ASTM 2009a and Section 10.2.5.1 in USEPA 2000). Overlying water should be a source of water that has been demonstrated to support survival, growth, and reproduction of *C. dilutus* (Section A.7.2.2 in ASTM 2009a and Section 15.2.2 in USEPA 2000). See additional guidance in ASTM (2009a,b) or in USEPA (2000) regarding demonstrating acceptable water quality for testing with mussels, mayflies, or oligochaetes.
3. Test organisms must be cultured and tested at the same temperature. Ideally, test organisms should be cultured in the same water that will be used in testing (Sections

A6.3.3 and A7.3.3 in ASTM 2009a and Sections 14.3.3.1 and 15.3.5.1 in USEPA 2000).

4. Before a sediment test is conducted in any new test facility, it is desirable to conduct a "non-toxicant" test, in which all test chambers contain a control sediment, and overlying water with no added test material. Survival, growth, or reproduction of the test organisms will demonstrate whether facilities, water, control sediment, and handling techniques are adequate to result in acceptable species-specific control numbers (Sections 9.1 and 11.14.1 in ASTM 2009a and Sections 6.1 and 9.3.3 in USEPA 2000). Before conducting tests with potentially contaminated sediment, it is strongly recommended that the laboratory conduct the tests with control sediment(s) alone. Results of these preliminary studies should be used to determine if the use of the control sediment and other test conditions (i.e., water quality) result in acceptable performance in the tests (Section 17.3 in ASTM 2009a and Section 17.4.3 in USEPA 2000).
5. The organisms should appear healthy, behave normally, feed well, and have low mortality in cultures, during holding (for example, <20% for 48 h before the start of a test), and in test controls (Section 11.4 in ASTM 2009a, Section 9.4 in USEPA 2000).
6. Intra-laboratory precision, expressed as a coefficient of variation, of the range for each type of test to be used in a laboratory can be determined by performing five or more tests with different batches of test organisms, using the same reference toxicant, at the same concentrations, with the same test conditions (for example, the same test duration, type of water, age of test organisms, feeding), and same data analysis methods (Sections 11.14.1 and 17.4.1 in ASTM 2009a and Sections 9.14.1 and 17.4.1 in USEPA 2000).
7. Laboratories should demonstrate that their personnel are able to recover an average of at least 90% of the organisms from whole sediment (Section 11.14.3 in ASTM 2009a and Section 9.14.3 in USEPA 2000).
8. Test organisms obtained from commercial sources should be shipped in well-oxygenated water in insulated containers to maintain temperature during shipment. Temperature and dissolved oxygen of the water in the shipping containers should be measured on arrival to determine if the organisms might have been subjected to low dissolved oxygen or temperature fluctuations (Section 12.2.3 in ASTM 2009a and Section 10.2.3 in USEPA 2000).
9. Organisms can be cultured using either static or renewal procedures. Renewal of water is recommended to limit loss of the culture organisms from a drop in dissolved oxygen or a buildup of waste products. In renewal systems, there should be at least one volume addition/day of culture water to each chamber (Section 12.2.5.2 in ASTM 2009a and Section 10.2.5.2 in USEPA 2000).
10. Numbers of bioaccumulation testing organisms (e.g., *Lumbriculus variegatus*) in a 4-day toxicity screening test should not be reduced significantly in the test sediment relative to the control sediment. Test organisms in bioaccumulation tests should burrow into test sediment. Avoidance of the test sediment by bioaccumulation testing organisms may decrease bioaccumulation (USEPA 2000, ASTM 2009d). Laboratories should monitor the frequency with which the populations of *L. variegatus* is doubling in the culture (the number of organisms) and record this information using control charts (the doubling rate would need to be estimated on a subset of animals from a mass culture). Records also should be kept on the frequency of restarting cultures. If static cultures are used, it may be desirable to measure water quality more frequently. Food used to culture organisms should be analyzed before the start of a test for compounds to be evaluated in

the bioaccumulation test. (USEPA 2000, ASTM 2009d).

A1.1 Detail Guidance

Specific sections dealing with laboratory performance in ASTM (2009a) for conducting sediment toxicity tests with the amphipod *Hyalella azteca* and the midge *Chironomus dilutus*. Analogous sections are included in USEPA (2000) for testing with *H. azteca* and *C. dilutus*. Analogous sections are included in ASTM (2009a,b,c) for testing with mussels, mayflies, and oligochaetes. Key recommendations are highlighted.

A1.2 Definition of Terms

3.1 The words "must", "should"," may", "can", and "might" have very specific meanings in this standard. "Must" is used to express an absolute requirement, that is, to state that a test ought to be designed to satisfy the specified conditions, unless the purpose of the test requires a different design. "Must" is used only in connection with the factors that relate directly to the acceptability of a test. "Should" is used to state that the specified condition is recommended and ought to be met if possible. Although the violation of one "should" is rarely a serious matter, violation of several will often render the results questionable. Terms such as "is desirable," "is often desirable," and "might be desirable" are used in connection with less important factors. "May" is used to mean "is (are) allowed to," "can" is used to mean "is (are) able to," and "might" is used to mean "could possibly." Thus, the classic distinction between "may" and "can" is preserved, and "might" is never used as a synonym for either "may" or "can."

A1.3 Summary of Standard

4.2.1.1.1 In general, the performance of test organisms in the negative control is used to judge the acceptability of a test, and either the negative control or reference sediment may be used to evaluate performance in the experimental treatments, depending on the purpose of the study. Any study in which organisms in the negative control do not meet performance criteria must be considered questionable because it suggests that adverse factors affected the response of test organisms. Key to avoiding this situation is using only control sediments that have a demonstrated record of performance using the same test procedure. This includes testing of new collections from sediment sources that have previously provided suitable control sediment.

A1.4 Facilities, Equipment, and Supplies

9.1 Before a sediment test is conducted in any new test facility, it is desirable to conduct a "non-toxicant" test, in which all test chambers contain a control sediment, and overlying water with no added test material (see 11.14). Survival, growth, or reproduction of the test organisms will demonstrate whether facilities, water, control sediment, and handling techniques are adequate to result in acceptable species-specific control numbers (for example, see 13.1.2). Evaluations may also be made on the magnitude of the within-chamber and between-chamber variance in a test. See 11.14.

A1.5 Quality Assurance and Quality Control

11.2 Performance-based Criteria:
11.2.1 The USEPA Environmental Monitoring Management Council (EMMC) recommended the use of performance-based methods in developing standards for chemical analytical methods **(89)**. Performance-based methods were defined by EMMC as a monitoring approach which permits the use of appropriate methods that meet pre-established demonstrated performance standards. Minimum required elements of performance, such as precision, reproducibility, bias, sensitivity, and detection limits should be specified and the method should be demonstrated to meet the performance standards.

11.2.2 In developing guidance for culturing test organisms to be included in this standard for sediment tests, it was generally agreed that no single method must be used to culture organisms. Success of a test relies on the health of the culture from which organisms are taken for testing. Having healthy organisms of known quality and age for testing is the key consideration relative to culture methods. Therefore, a performance-based criteria approach is the preferred method through which individual laboratories can evaluate culture health rather than requiring all laboratories to use the same culturing procedure. Performance-based criteria were chosen in USEPA **(2)** and in this standard to allow each laboratory to optimize culture methods while providing organisms that produce reliable and comparable test results. See 13.1.2, 14.1.2, and Annex A1 to Annex A7 for a listing of performance criteria for culturing and testing.

A1.6 Facilities, Equipment, and Test Chambers

11.3.3 Before a sediment test is conducted in a new facility, a "non-contaminant" test should be conducted in which all test chambers contain a control sediment and overlying water. This information is used to demonstrate that the facility, control sediment, water, and handling procedures provide acceptable responses of test organisms (see 11.14).

11.4 The organisms should appear healthy, behave normally, feed well, and have low mortality in cultures, during holding (for example, <20% for 48 h before the start of a test), and in test controls. Test organisms should be positively identified to species. Obtaining wild populations of organisms for testing should be avoided unless the ability of the wild population to cross-breed with existing laboratory populations has been determined (see 12.2.2).

11.5 The quality of water used for organism culturing and testing is extremely important. Overlying water used in testing and water used in culturing organisms should be uniform in quality. Acceptable water should allow satisfactory survival, growth, or reproduction of the test organisms. Test organisms should not show signs of disease or apparent stress.

A1.7 Quality of Test Organisms

11.8.1 It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (See Table 10 and Table 11 and Section 11.16). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity to select chemicals. The previous requirement for laboratories to conduct monthly reference-toxicity tests in an earlier version of this standard (Test Method E 1706-95b) has not been included as a requirement for testing sediments due to the inability of reference-toxicity tests to identify stressed populations of test organisms (McNulty *et al.* **(99)**). Physiological measurements such as lipid content might also provide useful information regarding the health of the cultures.

11.8.2 It is desirable to determine the sensitivity of test organisms obtained from an outside source. The supplier should provide data with the shipment describing the history of the sensitivity or organisms from the same source culture. The supplier should also certify the species identification of the test organisms, and provide the taxonomic references, or name(s) of the taxonomic expert(s) consulted.

11.8.3 All organisms in a test must be from the same source (Section 10.2.2). Organisms may be obtained from laboratory cultures or from commercial or government sources. The test organism used should be identified using an appropriate taxonomic key, and verification should be documented (e.g., **(165)**, Merritt and Cummins **(395)**). Obtaining organisms from wild populations should be avoided unless organisms are cultured through several generations in the laboratory. In addition, the ability of the wild population of sexually reproducing organisms to cross-breed with the existing laboratory population should be determined (Duan *et al.*, **(164)**). Sensitivity of the wild population to select chemicals (for example, Table 2) should also be documented.

A1.8 Demonstrating Acceptable Performance

11.14.1 Intralaboratory precision, expressed as a coefficient of variation, of the range for each type of test to be used in a laboratory can be determined by performing five or more tests with different batches of test organisms, using the same reference toxicant, at the same concentrations, with the same test conditions (for example, the same test duration, type of water, age of test organisms, feeding), and same data analysis methods. A reference-toxicant concentration series (0.5 or higher) should be selected that will provide partial mortalities at two or more concentrations of the test chemical (see Section 10.3.3). Information from previous tests can be used to improve the design of subsequent tests to optimize the dilution series selected for testing.

11.14.2 Before conducting tests with potentially contaminated sediment, it is strongly recommended that the laboratory conduct the tests with control sediment(s) alone. Results of these preliminary studies should be used to determine if the use of the control sediment and other test conditions (i.e., water quality) result in acceptable performance in the tests as outlined in Sections 13 and 14 and in Annex A1 to Annex A7.

11.14.3 Laboratories should demonstrate that their personnel are able to recover an average of at least 90% of the organisms from whole sediment. For example, test organisms could be added to control or test sediments, and recovery could be determined after 1 h **(161)**.

A1.9 Documenting Ongoing Laboratory Performance

11.15.2 A sediment test may be acceptable if specified conditions of a reference-toxicity test fall outside the expected ranges (see 11.10). Specifically, a sediment test should not be judged unacceptable if the LC50 for a given reference-toxicity test falls outside the expected range or if control survival in the reference-toxicity test is <90%. All the performance criteria outlined in 13.1.2 and 14.1.2 or in Annex A1 to Annex A7 must be considered when determining the acceptability of a sediment test. The acceptability of the sediment test would depend on the experience and judgment of the investigator and the regulatory authority.

A1.10 Collection, Culturing, and Maintaining Test Organisms

12.2.1 Acceptability of a culturing procedure is based in part on performance of organisms in culture and in the sediment test (see Section 5 and 11.2). No single technique for culturing test organisms is recommended. What may work well for one laboratory may not work as well for another laboratory. While a variety of culturing procedures are outlined in 12.3 for *H. azteca* and in 12.4 for *C. dilutus*, organisms must meet the test acceptability requirements listed in 13.1.2 or 14.1.2. Culturing procedures are outlined for the additional test organisms in Annex A1 to Annex A7.

12.2.2 All organisms in a test must be from the same source. Organisms may be obtained from laboratory cultures or from commercial or government sources (see Table 12). The test organism used should be identified using an appropriate taxonomic key and verification should be documented. Obtaining organisms from wild populations should be avoided unless organisms are cultured through several generations in the laboratory. In addition, the ability of the wild population to crossbreed with existing laboratory populations should be determined. The sensitivity of the wild population to select chemicals (for example, Table 2) should also be documented (Duan *et al.* **(164)**).

12.2.3 Test organisms obtained from commercial sources should be shipped in well-oxygenated water in insulated containers to maintain temperature during shipment. Temperature and dissolved oxygen of the water in the shipping containers should be measured on arrival to determine if the organisms might have been subjected to low dissolved oxygen or temperature fluctuations. The temperature of the shipped water should be gradually adjusted to the desired culture temperature at a rate not exceeding 2°C/24 h. Additional reference-toxicity testing is suggested if organisms are not cultured at the testing laboratory (see 11.16).

12.2.4 A group of organisms should not be used for a test if they appear to be unhealthy, discolored, or otherwise stressed (for example, >20% mortality for 48 h before the start of a test). If the organisms fail to meet these criteria, the entire batch should be discarded and a new batch should be obtained. All organisms should be as uniform as possible in age and life stage. Test organisms should be handled as little as possible. When handling is necessary, it should be done as gently, carefully, and as quickly as possible.

12.2.5 *Hyalella azteca, C. dilutus* , and the test organisms described in Annex A1 to Annex A7 can be cultured in a variety of waters. Water of a quality sufficient to culture fathead minnows (*Pimephales promelas*) or cladocerans will generally be adequate.

12.2.5.1 Variable success has been reported using reconstituted waters to culture or test *H. azteca* (See 7.1.3).

12.2.5.2 Organisms can be cultured using either static or renewal procedures. Renewal of water is recommended to limit loss of the culture organisms from a drop in dissolved oxygen or a buildup of waste products. In renewal systems, there should be at least one volume addition/day of culture water to each chamber. In static systems, the overlying water volume should be changed at least weekly by siphoning down to a level just above the substrate and slowly adding fresh water. Extra care should be taken to ensure that proper water quality is maintained in static systems. For example, aeration is needed in static systems to maintain dissolved oxygen at >2.5 mg/L.

A1.11 Demonstrating Acceptable Laboratory Performance

17.4.1 Intralaboratory precision, expressed as a coefficient of variation (CV), of the range for each type of test to be used in a laboratory can be determined by performing five or more tests with different batches of test organisms, using the same reference toxicant, at the same concentrations, with the same test conditions (for example, the same test duration, type of water, age of test organisms, feeding), and same data analysis methods. A reference-toxicant concentration series (0.5 or higher) should be selected that will consistently provide partial mortalities at two or more concentrations of the test chemical (see 11.14, Table 10 and Table 11). See Section 11.16 for additional detail regarding reference-toxicity testing.

17.4.2 It is desirable to determine the sensitivity of test organisms obtained from an outside source. The supplier should provide data with the shipment describing the history of the sensitivity of organisms from the same source culture.

17.4.3 Before conducting tests with potentially contaminated sediment, it is strongly recommended that the laboratory conduct the tests with control sediment(s) alone. Results of these preliminary studies should be used to determine if the use of the control sediment and other test conditions (i.e., water quality) result in acceptable performance in the tests as outlined in Sections 13 and 14 and in Annex A1 to Annex A7.

17.4.4 A control chart can be prepared for each combination of reference toxicant and test organism. Each control chart should include the most current data. Endpoints from five tests are adequate for establishing the control charts. In this technique, a running plot is maintained for the values (X_i) from successive tests with a given reference toxicant (see Fig. 16), and the endpoint (LC50, NOEC, ICp) are examined to determine if they are within prescribed limits. Control charts as described in **(3, 157)** are used to evaluate the cumulative trend of results from a series of samples. The mean and upper and lower control limits (±2 SD) are recalculated with each successive test result.

A1.12 Procedure for Conducting a *Hyalella azteca* 42-day Test for Measuring Effects of Sediment-associated contaminants on survival, growth, and reproduction

A6.1.3 The procedure outlined in Section A6.2 is based on procedures described in Ingersoll *et al.* **(82)**. The sediment exposure starts with 7- to 8-day-old amphipods. On Day 28, amphipods are isolated from the sediment and placed in water-only chambers where reproduction is measured on Day 35 and 42. Typically, amphipods are first in amplexus at about Day 21 to 28 with release of the first brood between Day 28 to 42. Endpoints measured include survival (Day 28, 35 and 42), growth (as length or dry weight measured on Day 28 and 42), and reproduction (number of young/female produced from Day 28 to 42). The procedures described in Section A6.2 include measurement of a variety of lethal and sublethal endpoints; minor modifications of the basic methods can be used in cases where only a subset of these endpoints is of interest.

A6.2.2 The 42-day sediment toxicity test with *H. azteca* is conducted at 23°C with a 16L:8D photoperiod at an luminance of about 100 to 1000 lux (Table A6.1). Test chambers are 300-mL high-form lipless beakers containing 100 mL of sediment and 175 mL of overlying water. Amphipods in each test chamber are fed 1.0 mL of YCT daily (Annex A8). Each test chamber receives 2 volume additions/day of overlying water. Water renewals may be manual or automated. Zumwalt *et al.* **(130)**, Benoit *et al.*, **(129)** and USEPA **(2)** describe water-renewal systems that can be used to deliver overlying water. Overlying water should be a source of water that has been demonstrated to support survival, growth, and reproduction of *H. azteca* in culture. McNulty *et al.* **(366)** and Kemble *et al.* **(61)** observed poor survival of *H. azteca* in tests conducted 14 to 28 day using a variety of reconstituted waters including reconstituted water (reformulated moderately hard reconstituted water) described in Smith *et al.* **(112)** and described in an earlier version of this test method (Test Method E 1706-95b). Borgmann **(367)** described a reconstituted water that was used successfully to maintain *H. azteca* in culture; however, some laboratories have not had success when using this reconstituted water in the 42-day test (T.J. Norberg-King, USEPA, Duluth, MN, personal communication). For site-specific evaluations, the characteristics of the overlying water should be as similar as possible to the site where sediment is collected. Requirements for test acceptability are summarized in Table A6.3.

A6.3.3 Acclimation: Test organisms must be cultured and tested at the same temperature. Ideally, test organisms should be cultured in the same water that will be used in testing. However, acclimation of test organisms to the test water is not required. If test organisms are to be acclimated, they could be held for 2 h in a 50 to 50 mixture of culture water to overlying water, then for 2 h in a 25 to 75 mixture of culture water to overlying water, followed by a transfer into 100% overlying water for 2 h **(103)**.

A1.13 Procedure for Conducting a life-cycle test for measuring effects of sediment associated contaminants on *Chironomus dilutus*

A7.1.3 The long-term sediment toxicity test with the midge, *Chironomus dilutus*, is a life-cycle test in which the effects of sediment exposure on survival, growth, emergence, and reproduction are assessed (Benoit *et al.* **(70)**). Procedures for conducting the long-term test with *C. dilutus* are described in Section A7.2. The test is started with newly hatched larvae (<24-h old) and continue through emergence, reproduction, and hatching of the F_1 generation. Survival is determined at 20 days and at the end of the test (about 50 to 65 days). Growth is determined at 20 day, which corresponds to the 10-day endpoint in the 10-day *C. dilutus* growth test started with 10-day old larvae (Section 14). From Day 23 to the end of the test, emergence and reproduction are monitored daily. The number of eggs is determined for each egg case, which is incubated for 6 days to determine hatching success. Each treatment of the life-cycle test is ended separately when no additional emergence has been recorded for 7 consecutive days (the 7-day criterion). When no emergence is recorded from a treatment, ending of that treatment should be based on the control sediment using this 7-day criterion. Table 6.1 and Section A7.5 outline equipment and supplies needed to conduct this test. The procedures described in Section A7.2 include measurement of a variety of lethal and sublethal endpoints; minor modifications of the basic methods can be used in cases where only a subset of these endpoints is of interest.

A7.2.2 The long-term sediment toxicity test with *C. dilutus* is conducted at 23°C with a 16L:8D photoperiod at an illuminance of about 100 to 1000 lux (Table A7.1). Test chambers are 300-mL high-form lipless beakers containing 100 mL of sediment and 175 mL of overlying water. Each test chamber receives 2 volume additions/day of overlying water. Water renewals may be manual or automated. Zumwalt *et al.* **(130)**. Benoit *et al.* **(129)** and USEPA **(2)** describe water-renewal systems that can be used to deliver overlying water. Overlying water should be a source of water that has been demonstrated to support survival, growth, and reproduction of *C. dilutus* in culture. For site-specific evaluations, the characteristics of the overlying water should be as similar as possible to the site where sediment is collected. Requirements for test acceptability are summarized in Table A7.3.

A1.14 Report

16.1 The record of the results of an acceptable sediment test should include the following information either directly or by referencing available documents:

16.1.1 Name of test and investigator(s), name and location of laboratory, and dates of start and end of test.

16.1.2 Source of control or test sediment, method for collection, handling, shipping, storage, and disposal of sediment.

16.1.3 Source of test material, lot number if applicable, composition (identities and concentrations of major ingredients and impurities if known), known chemical and physical properties, and the identity and concentration(s) of any solvent used.

16.1.4 Source and characteristics of overlying water, description of any pretreatment, and results of any demonstration of the ability of an organism to survive or grow in the water.

16.1.5 Source, history, and age of test organisms; source, history, and age of brood stock, culture procedures; and source and date of collection of the test organisms, scientific name, name of person who identified the organisms and the taxonomic key used, age or life stage, means and ranges of weight or length, observed diseases or unusual appearance, treatments, holding, and acclimation procedures.

16.1.6 Source and composition of food, concentrations of test material and other contaminants, procedure used to prepare food, feeding methods, frequency, and ration.

16.1.7 Description of the experimental design and test chambers, the depth and volume of sediment and overlying water in the chambers, lighting, number of test chambers and number of test organisms/treatment, date and time test starts and ends, temperature measurements, dissolved oxygen concentration (as percent saturation), and any aeration used before starting a test and during the conduct of a test.

16.1.8 Methods used for physical and chemical characterization of sediment.

16.1.9 Definition(s) of the effects used to calculate LC50 or EC50s, biological endpoints for tests, and a summary of general observations of other effects.

16.1.10 Methods used for statistical analyses of data: (*1*) summary statistics of the transformed or raw data as applicable (for example, mean, standard deviation, coefficient of variation, precision and bias); (*2*) hypothesis testing (raw data, transformed data, null hypothesis, alternate hypothesis, target Type I and II error rates, statistics used (including calculation of test statistic)), decision rule used (for example, *W* statistic >0.65 results in the rejection of the null hypothesis), calculated test statistic and decision rule result, achieved Type I and II error rates (for some discrete tests, achieved error rates only approximate the target rates); *(3)* results of regression analyses (parameters of regression fit, uncertainty limits on the regression parameters, correlation coefficient).

16.1.11 Summary of general observations on other effects or symptoms.

16.1.12 Anything unusual about the test, any deviation from these procedures, and any other relevant information.

16.2 Published reports should contain enough information to clearly identify the methodology used and the quality of the results.

Appendix A, Table 1. Test Conditions for Conducting a Long-term Sediment Toxicity Test (adapted from Tables A6.1 and A7.1 in ASTM 2009a [analogous tables are included in USEPA 2000]). See additional detail outlined in ASTM (2008d) and in Ingersoll *et al.* (2008) for testing mussels. See ASTM (2009a) for additional guidance on testing with mayflies and see ASTM (2009a) or USEPA (2000) for additional guidance on conducting 10-d sediment toxicity tests with freshwater invertebrates.

Parameter	Conditions
1. Test type	Sediment toxicity test with renewal of overlying water conducted with control sediment and field-collected sediment samples. Species to be tested: the amphipod *Hyalella azteca* the midge *Chironomus dilutus*, juvenile mussels (fatmucket, *Lampsilis siliquoidea* or rainbow mussels, *Villosa iris*), and mayflies (Hexagenia sp.; likely a mixture of *H. limbata* and *H. rigida*)
2. Temperature	23 ± 1°C
3. Light quality	Wide-spectrum fluorescent lights
4. Illuminance	About 200 lux
5. Photoperiod	16L:8D
6. Test chamber	300-mL high-form lipless beaker (1-L beakers for testing with mayflies)
7. Sediment volume	100 mL (sediments sieved to <2 mm at the collection site) and added to exposure beakers on about Day-7; sediments in mussel test sieved to <0.25 mm). 200 mL sediment for testing with mayflies
8. Overlying water volume	175 mL (700 mL for testing with mayflies)
9. Renewal of overlying water	2 volume additions/d
10. Age of organisms	Amphipods and midge: About 7-d-old organisms Mussels: 2- to 4-month old juvenile mussels (Wang *et al.* 2007, Ingersoll *et al.* 2008) Mayflies: 6- to 8-weeks old (Day *et al.* 1995, 1998) Archive 20 amphipods on Day 0 for length measurement and 4 replicates of 10 midge for starting ash-free dry weight and 10 mayflies for starting dry weight; photograph mussels before placement into each replicate on Day 0
11. Number of organisms/chamber	Amphipods, mussels, mayflies: 10 Midge: 12
12. Number of toxicity replicate chambers/treatment	Amphipods: 12 (4 sampled on Day 28 for survival and growth and 8 on Day 42 for survival, growth, and reproduction). Two additional chemistry replicate/treatment (for SEM/AVS/peeper metal sampling at the beginning and end of a test) Midge: 16 (4 sampled on Day 13 for survival and growth, 8 for emergence and reproduction, and 4 for auxiliary males [started late on Day 2]). Two additional chemistry replicate/treatment (for SEM/AVS/peeper metal sampling) Mussels and Mayflies: 4 (sampled on Day 28 for survival and growth). Two additional chemistry replicate/treatment for SEM/AVS/peeper metal sampling at the beginning and end of a test)
13. Feeding	Amphipods: YCT food, fed 1.0 mL (1800 mg/L stock) daily to each test chamber Midge: Tetrafin® goldfish food, fed 1.5 mL daily to each test chamber (1.5 mL contains 6.0 mg of dry solids) Mussels: Non-viable algae daily (Wang *et al.* 2007, ASTM 2008d) Mayflies: YCT food, fed 4.0 mL (800 mg/L stock) daily to each test chamber (Day *et al.* 1995, 1998).
14. Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L
15. Overlying water	Well water diluted with deionized water to a hardness of about 100 mg/L (as $CaCO_3$), alkalinity 85 mg/L (as $CaCO_3$), and pH about 8.0. Note: these

Parameter	Conditions
	conditions may be modified depending on the testing laboratory. Importantly, the overlying water must have at least 10 mg Cl/L.
16. Test chamber cleaning	If screens become clogged during a test; gently brush the outside of the screen
17. Overlying water quality	Hardness, alkalinity, pH and ammonia at the beginning and end of a test and on the days that growth is sub-sampled. Temperature daily. Dissolved oxygen (DO) and pH three times/week. Concentrations of DO should be measured more often if DO has declined by more than 1 mg/L since previous measurement.
18. Chemistry sampling of pore water and sediment	Pore water (sampled on about Day -7): centrifugation at 4ºC for 15 minutes at 5,200 rpm; 0.5 L sediment/treatment): Ammonia, pH, free sulfide, hardness, alkalinity, conductivity, dissolved oxygen, major cations and anions, dissolved organic carbon. Sediment (sampled on about Day -7): (1) PCB Aroclors, homologs and congeners, (2) TAL metals, (3) OCl pesticides, (4) PAHs, and (5) PCDDs/PCDFs. Note: this listing of organic analyses needs to be evaluated based on historic data for these compounds. Sediment (sampled during exposures): Simultaneously extracted metals, acid volatile sulfide, peeper pore-water metals (beginning and end of each toxicity test in separate chemistry beakers)
19. Endpoints	Amphipods: 28-d survival and growth (length, weight, total biomass), 35-d reproduction, and 42-d survival, growth, reproduction, and sensitivity of the F1 offspring with a NaCl challenge toxicity test. Midge: 13-d survival and ash-free-dry weight, female and male emergence, adult mortality, the number of egg cases oviposited, the number of eggs produced, the number of hatched eggs, and sensitivity of the F1 offspring with a NaCl challenge toxicity test. Sensitivity of offspring will be determined by estimation of an LT50 value or LC50 value where adequate numbers of animals are available. Mussels and Mayflies: 28-d survival and growth (shell length, weight, total biomass).
20. Test acceptability	Amphipods: Minimum mean control survival of 80% on Day 28. Additional performance-based criteria specifications are outlined in Table 3. Midge: No test acceptability requirements have been established for long-term tests starting with 7-d-old larvae. Average survival of *C. dilutus* in the control sediment should be greater than or equal to 70% at Day 20 and greater than or equal to 65% at the end of the test. Emergence should be greater than or equal to 50%. Mussels: Minimum mean control survival of 80% on Day 28 (ASTM 2008d, Ingersoll *et al*. 2008). Mayflies: Minimum mean control survival of 80% on Day 28 (ASTM 2008a). Additional general performance-based criteria specifications outlined in Table 3 are also applicable to mussels and mayflies.

Appendix A, Table 2. General Activity Schedule for Conducting a Long-term Sediment Toxicity Test with *Hyalella azteca* or with Mussels (adapted from Table A6.2 in ASTM 2009a, an analogous table is included in USEPA 2000). See ASTM (2009a) for additional guidance on testing with mayflies.

Day	Activity
Pre-Test	
–7	Sample sediments for physical and chemical characteristics and sample pore water by centrifugation for water quality analyses. Place sediments into exposure beakers for about a 7-d equilibration period before the start of the toxicity tests. Add sediment into each test chamber, place chambers into exposure system, and add overlying water.
–2	Isolate amphipods or mussels from culture and feed and observe isolated amphipods or mussels to evaluate health.
-1	1. Check cultures for hatch and development. 2. Start water renewal.
Sediment Test	
0	Measure total water quality of overlying water (pH, temperature, dissolved oxygen, hardness, alkalinity, conductivity, ammonia). Transfer ten amphipods or mussels into each test chamber. Release organisms under the surface of the water. Add appropriate food to each test chamber. Archive 20 amphipods for length measurement and photograph mussels to be placed in each replicate exposure chamber. Observe behavior of test organisms. Place peepers in sediment chemistry beakers (and sample AVS, SEM, and metals in peeper samples on Day 7).
7	Sample metals in peeper samples, and AVS and SEM.
1 to 27	Feed test organisms. Measure temperature daily, and dissolved oxygen (DO) and pH three times/week. Observe behavior of test organisms.
14 to 21	Place peepers in chemistry beakers on Day 14 and sample peepers from chemistry beakers on Day 21.
21	Place peepers in sediment chemistry beakers
28	Measure temperature, dissolved oxygen, pH, hardness, alkalinity, conductivity and ammonia. End the sediment-exposure portion of the test by collecting the test organisms with a #40 mesh sieve (425-µm mesh; U.S. standard size sieve). Sample metals in peeper samples, and AVS and SEM. Use four amphipod replicates and four mussel replicates for growth measurements: count survivors and preserve organisms in sucrose-formalin solution for growth measurements. Eight amphipod replicate beakers for reproduction measurements: Place survivors in individual replicate water-only beakers containing 5-ml of clean sand and add food to each test beaker/d and 2 volume additions/d of overlying water.
Amphipod Reproduction Phase	
29 to 35	Feed daily. Measure temperature daily, conductivity weekly, DO and pH three times a week. Measure hardness and alkalinity weekly. Observe behavior of test organisms.
35	Record the number of surviving adults and remove offspring. Return adults to their original individual beakers and add food. Conduct NaCl challenge test with offspring (depending on the number of offspring available).
36 to 41	Feed daily. Measure temperature daily, conductivity weekly, DO and pH three times a week. Measure hardness and alkalinity weekly. Observe behavior of test organisms.
41	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, conductivity, ammonia).
42	Record the number of surviving adults and offspring. Surviving adult amphipods on Day 42 are preserved in sucrose-formalin solution. The number of adult males in each beaker is determined from this archived sample. This information is used to calculate the number of young produced per female per replicate from Day 28 to Day 42. Conduct NaCl challenge test with offspring (depending on the number of offspring available).

Appendix A, Table 3. General Activity Schedule for Conducting a Long-term Sediment Toxicity Test with *Chironomus dilutus* (adapted from Table A7.2 in ASTM 2009a, an analogous table is included in USEPA 2000 with 7-d-old larvae used to start the exposures).

Day	Activity
–-11	Start reproduction flask with cultured adults (1:3 male: female ratio). For example for 15 to 25 egg cases, 10 males and 30 females are typically collected. Egg cases typically range from 600 to 1500 egg/case.
–10	Collect egg cases (a minimum of 6 to 8) and incubate at 23°C.
–9	Check egg cases for viability and development.
–7	Sample sediments for physical and chemical characteristics and sample pore water by centrifugation for water quality analyses. Place sediments into exposure beakers for about a 7-d equilibration period before the start of the toxicity tests. Add sediment into each test chamber, place chambers into exposure system, and add overlying water.
–-1	1. Check cultures for hatch and development.
	2. Start water renewal.
0	1. Transfer larvae into exposure chambers by transferring larvae still in their culture tubes into the exposure chambers. Add 1.5 mL food to each test beaker with sediment before the larvae are added. Add 12 larvae to each replicate beaker (beakers are chosen by random block assignment).
	2. Measure temperature, pH, hardness, alkalinity, dissolved oxygen, conductivity and ammonia at start of test, and on day 13. 3. Place peepers in sediment chemistry beakers.
1–End	On a daily basis, add 1.5 mL food to each beaker. Measure temperature daily. Measure the pH and dissolved oxygen three times a week during the test. If the DO has declined more than 1 mg/L since previous reading, increase frequency of DO measurements and aerate if DO continues to be less than 2.5 mg/L.
-1	For auxiliary male production, start reproduction flask with culture adults (e.g., 10 males and 30 females; 1:3 male to female ratio).
0-2	Set up schedule for auxiliary male beakers (4 replicates/treatment) same as that described above for the beginning of the test (note on Day 2 because Day 3 would be a Saturday).
7	Sample metals in peeper samples, and AVS and SEM
12	In preparation for weight determinations, ash weigh-pans at 550 °C for 2 h. Note that the weigh boats should be ashed before use to eliminate weighing errors due to the pan oxidizing during ashing of samples.
13	Randomly select four replicates from each treatment and sieve the sediment to recover larvae for growth and survival determinations. Pool all living larvae per replicate and dry the sample to a constant weight (e.g., 60°C for 24 h). Install emergence traps on each reproductive replicate beaker. Measure overlying water quality.
14	The sample with dried larvae is brought to room temperature in a desiccator and weighed to the nearest 0.01 mg. The dried larvae in the pan are then ashed at 550°C for 2 h. The pan with the ashed larvae is then re-weighed and the tissue mass of the larvae determined as the difference between the weight of the dried larvae plus pan and the weight of the ashed larvae plus pan.
16–End	On a daily basis, record emergence of males and females, pupal, and adult mortality, and time to death for previously collected adults. Each day, transfer adults from each replicate to a corresponding reproduction/oviposition (R/O) chamber. Transfer each primary egg case from the R/O chamber to a corresponding Petri dish to monitor incubation and hatch. Record each egg case oviposited, number of eggs produced (using either the ring or direct count methods), and number or hatched eggs. If it is difficult to estimate the number of eggs in an egg case, use a direct count to determine the number of eggs; however the hatchability data will not be obtained for this egg case. Conduct NaCl challenge test with offspring (depending on the number of offspring available).
14 to 21	Place peepers in chemistry beakers on Day 14 and sample peepers from chemistry beakers on Day 21.
21	Place peepers in sediment chemistry beakers
23	Place emergence traps on auxiliary male replicate beakers.

Day	Activity
27–End	Transfer males emerging from the auxiliary male replicates to individual inverted Petri dishes. The auxiliary males are used for mating with females from corresponding treatments from which most of the males had already emerged or in which no males emerged.
28	Sample metals in peeper samples, and AVS and SEM,
33–End	After 7 d of no recorded emergence in a given treatment, end the treatment by sieving the sediment to recover larvae, pupae, or pupal exuviae. When no emergence occurs in a test treatment, that treatment can be ended once emergence in the control sediment has ended using the 7-day criterion.

Appendix A, Table 4. Test Acceptability Requirements for a Long-term Sediment Toxicity Test with *Hyalella azteca* (adapted from Table A6.3 in ASTM 2009a, an analogous table is included in USEPA 2000). See ASTM (2009a,b) for additional guidance on testing with mayflies and mollusks.

A.	It is recommended for conducting the 42-day test with *H. azteca* that the following performance criteria be met:
	1. Age of *H. azteca* at the start of the test should be 7- to 8-day old. Starting a test with substantially younger or older organisms may compromise the reproductive endpoint.
	2. Average survival of *H. azteca* in the control sediment on Day 28 should be greater than or equal to 80%.
	3. Laboratories participating in round-robin testing (Section 17.6 of ASTM 2009a) reported after 28-day sediment exposures in a control sediment (West Bearskin), survival >80% for >88% of the laboratories; length >3.2 mm/individual for >71% of the laboratories; and dry weight >0.15 mg/individual for 66% of the laboratories. Reproduction from Day 28 to Day 42 was >2 young/female for 71% of the laboratories participating in the round-robin testing. Reproduction was more variable within and among laboratories; hence, more replicates might be needed to establish statistical differences among treatments with this endpoint.
	4. Hardness, alkalinity, and ammonia in the overlying water typically should not vary by more than 50% during the sediment exposure, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.
B.	Performance-based criteria for culturing *H. azteca* include the following:
	1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 11.16.2 of ASTM 2009a). Data from these reference toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
	2. Laboratories should track parental survival in the cultures and record this information using control charts if known-age cultures are maintained. Records should also be kept on the frequency of restarting cultures and the age of brood organisms.
	3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature in the cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.
	4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
	5. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.
C.	Additional requirements:
	1. All organisms in a test must be from the same source.
	2. Storage of sediments collected from the field should follow guidance outlined in Section 10.2 of ASTM (2009a). Sediment toxicity tests will be started within 2 months of collection from the field.
	3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
	4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
	5. Test organisms must be cultured and tested at 23°C (±1 °C).
	6. The mean of the daily test temperature must be within ± 1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.
	7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Appendix A, Table 5. Test Acceptability Requirements for a Long-term Sediment Toxicity Test with *Chironomus dilutus* (adapted from Table A7.3 in ASTM 2009a, an analogous table is included in USEPA 2000). Note the requirements below are for tests started with <24-h old larvae. It is anticipated that similar requirements will be met for the test starting with 7-d-old larvae.

A.	It is recommended for conducting a long-term test with C. dilutus that the following performance criteria be met:
	1. Tests must be started with less than 1-day (<24 h) old larvae. Starting a test with substantially older organisms may compromise the emergence and reproductive endpoint.
	2. Average survival of *C. dilutus* in the control sediment should be greater than or equal to 70% at Day 20 and greater than or equal to 65% at the end of the test.
	3. Average size of *C. dilutus* in the control sediment at 20 d must be at least 0.6 mg/surviving organism as dry weights or 0.48 mg/surviving organism as AFDW. Emergence should be greater than or equal to 50%. Experience has shown that pupae survival is typically >83% and adult survival is >96%. Time to death after emergence is <6.5 d for males and <5.1 d for females. The mean number of eggs/egg case should be greater than or equal to 800 and the percent hatch should be greater than or equal to 80%. See Sections A7.1.3 and 17.6 in ASTM (2009a) for a summary of performance in round robin testing.
	4. Hardness, alkalinity and ammonia in the overlying water within a treatment typically should not vary by more than 50% during the test and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.
B.	Performance-based criteria for culturing *C. dilutus* include the following:
	1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 11.16.2 of ASTM 2009a). Data from these reference toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
	2. Laboratories should keep a record of time to first emergence for each culture and record this information using control charts. Records should also be kept on the frequency of restarting cultures.
	3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature in the cultures should be recorded dally. If static cultures are used, it may be desirable to measure water quality more frequently.
	4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
	5. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.
C.	Additional requirements:
	1. All organisms in a test must be from the same source.
	2. Storage of sediments collected from the field should follow guidance outlined in Section 10.2 of ASTM (2009a). Sediment toxicity tests will be started within 2 months of collection from the field.
	3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
	4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
	5. Test organisms must be cultured and tested at 23°C (+1 °C).
	6. The mean of the daily test temperature must be within ± 1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.
	7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Appendix A, Table 6. Recommended Test Conditions for Conducting a 28-Day Sediment Bioaccumulation Test with *Lumbriculus variegatus* (adapted from Table A8.3 in ASTM 2009c, an analogous table is included in USEPA 2000).

Parameter	Conditions
1. Test type	Sediment bioaccumulation test with renewal of overlying water
2. Temperature	23°C
3. Light quality	Wide-spectrum fluorescent lights
4. Illuminance	100 to 1000 lux
5. Photoperiod	16L:8D
6. Test chamber	2-L beaker
7. Sediment volume	600 mL
8. Overlying water volume	1,400 mL
9. Renewal of overlying water	80% volume exchange three times weekly
10. Age of test organisms	Adults
11. Loading of organisms in chamber	1g wet weight per replicate
12. Number of replicate chambers/treatment	Five replicates are recommended for routine testing. See Table B2 in Attachment B for a total number of replicate beakers to be tested.
13. Feeding	None
14. Aeration	Trickle-flow (1-2 bubbles per second)
15. Overlying water	Well water diluted with deionized water to a hardness of about 100 mg/L (as $CaCO_3$), alkalinity 85 mg/L (as $CaCO_3$), and pH about 8.0.
16. Test chamber cleaning	Not necessary.
17. Overlying water quality	Hardness, alkalinity, pH, and ammonia at the beginning and end of a test. Temperature daily.
18. Chemistry sampling of pore water and sediment	Pore water (sampled on about Day -7): centrifugation at 4ºC for 15 minutes at 5,200 rpm; 0.5 L sediment/treatment): Ammonia, pH, free sulfide, hardness, alkalinity, dissolved oxygen. Sediment (sampled on about Day -7): (1) PCB Aroclors, homologs and congeners, (2) TAL metals, (3) OCl pesticides, (4) PAHs, and (5) PCDD/DFs. Note: see comments on Table 1 regarding measurement of organic contaminants of interest. Sediment (sampled during exposures): Simultaneously extracted metals, acid volatile sulfide, peeper pore-water metals (beginning and end of each bioaccumulation test in separate chemistry beakers)
19. Test duration	28 days
20. Endpoint	Bioaccumulation
21. Test acceptability	Performance-based criteria specifications outlined in Table 9.

Appendix A, Table 7. General Activity Schedule for Conducting a 28-Day Sediment Bioaccumulation Test with *Lumbriculus variegatus* (adapted from Table A8.4 in ASTM 2009c, an analogous table is included in USEPA 2000).

Day	Activity
–7	Sample sediments for physical and chemical characteristics and sample pore water for water quality analyses (associated with amphipod and midge testing, see Table 1). Place sediments into exposure beakers for about a 7-d equilibration period before the start of the bioaccumulation test. Add sediment into each test chamber, place chambers into exposure system, and start renewing overlying water.
–1	Isolate worms for conducting bioaccumulation test..
0	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, and ammonia). Sample a subset of worms used to start the test for residue analyses. Transfer appropriate amount of worms (based on weight) into each test chamber. Observe the behavior of test organisms. Place peepers in sediment chemistry beakers.
7	Sample metals in peeper samples, and AVS and SEM
1-27	Measure temperature and dissolved oxygen. Observe the behavior of test organisms.
21	Place peepers in sediment chemistry beakers.
28	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, and ammonia). End the uptake by collecting the worms with a sieve. Separate any indigenous organisms from *L. variegatus*. Determine the weight of survivors. Eliminate the gut contents of surviving worms in water for 6 to 8 h. Longer purging periods (not to exceed 24 h) may be used if all target analytes have Log K_{ow} >5 (see A8.4.7.3).Sample metals in peeper samples, and AVS and SEM .

Appendix A, Table 8. Test Acceptability Requirements for a 28-Day Sediment Bioaccumulation Test with *Lumbriculus variegatus* (adapted from Table A8.4 in ASTM 2009c, an analogous table is included in USEPA 2000).

A.		It is recommended for conducting a 28-day test with *L. variegatus* that the following performance criteria are met:
	1.	Numbers of *L. variegatus* in a 4-day toxicity screening test should not be reduced significantly in the test sediment relative to the control sediment.
	2.	Test organisms should burrow into test sediment. Avoidance of the test sediment by *L. variegatus* may decrease bioaccumulation.
	3.	The hardness, alkalinity, pH, and ammonia of overlying water within a treatment typically should not vary by more than 50 % during the test and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.
B.		Performance-based criteria for culturing *L. variegatus* include the following:
	1.	It may be desirable for laboratories to perform periodically 96-h water-only reference toxicity tests to assess the sensitivity of culture organisms (see ASTM 2009a). Data from these reference toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
	2.	Laboratories should monitor the frequency with which the population is doubling in the culture (the number of organisms) and record this information using control charts (the doubling rate would need to be estimated on a subset of animals from a mass culture). Records also should be kept on the frequency of restarting cultures. If static cultures are used, it may be desirable to measure water quality more frequently.
	3.	Food used to culture organisms should be analyzed before the start of a test for compounds to be evaluated in the bioaccumulation test.
	4.	Laboratories should record the following water quality characteristics of the cultures at least quarterly and the day before the start of a sediment test: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperatures of the cultures should be recorded daily.
	5.	Laboratories should characterize and monitor the background contamination and nutrient

		quality of food if problems are observed in culturing or testing organisms. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.
C.		Additional requirements:
	1.	All organisms in a test must be from the same source.
	2.	Storage of sediment collected from the field should follow guidance outlined in Section 10.7 of ASTM (2009c).
	3.	All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water
	4.	Negative-control sediment or appropriate solvent controls, must be included in a test. The concentration of solvent used must not affect test organisms adversely.
	5.	Culture and test temperatures must be the same. Acclimation of test organisms to the test water is not required.
	6.	The daily mean test temperature must be within ±1°C of the desired temperature. The instantaneous temperature must always be within ±3°C of the desired temperature.
	7.	Natural physicochemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.